

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年6月30日(30.06.2022)



(10) 国際公開番号

WO 2022/137778 A1

(51) 国際特許分類:
G06Q 10/06 (2012.01)

(21) 国際出願番号: PCT/JP2021/039367

(22) 国際出願日: 2021年10月25日(25.10.2021)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2020-212788 2020年12月22日(22.12.2020) JP

(71) 出願人: 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).

(72) 発明者: 野澤 拓磨 (NOZAWA, Takuma); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 小山田 昌史 (OYAMADA, Masafumi); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 董于洋 (DONG, Yuyang); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 草野 元紀 (KUSANO, Genki);

〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).

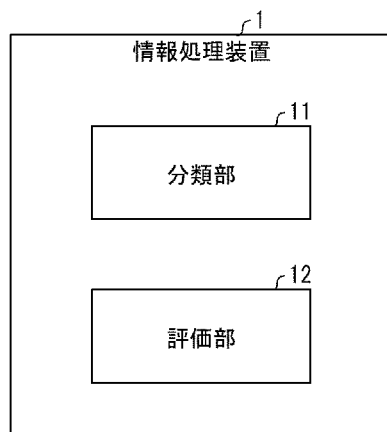
(74) 代理人: 特許業務法人 H A R A K E N Z O W O R L D P A T E N T & T R A D E M A R K (HARAKENZO WORLD PATENT & TRADEMARK); 〒5300041 大阪府大阪市北区天神橋2丁目北2番6号 大和南森町ビル Osaka (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) Title: INFORMATION PROCESSING DEVICE, ANALYSIS METHOD, AND ANALYSIS PROGRAM

(54) 発明の名称: 情報処理装置、分析方法、および分析プログラム

[図1]



- 1 Information processing device
11 Classification unit
12 Evaluation unit

(57) Abstract: The present invention enables insights between a plurality of datasets to be detected. An information processing device (1) comprises: a classification unit (11) for grouping, for each insight to be detected, insight subjects which are the data generated from each of a plurality of datasets by associating a plurality of data items included in the dataset; and an evaluation unit (12) for calculating an evaluation value for assessing the presence of insights with regard to a combination of the plurality of grouped insight subjects.

(57) 要約: 複数のデータセット間におけるインサイトの検出を可能にする。情報処理装置(1)は、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する分類部(11)と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する評価部(12)と、を備える。

[続葉有]

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

明 細 書

発明の名称： 情報処理装置、分析方法、および分析プログラム
技術分野

[0001] 本発明は、データセットの解析を行う情報処理装置等に関する。

背景技術

[0002] 近年、様々な分野において、データを収集し、そのデータを分析することにより、人にとって意味のある知見を見出すことが行われている。このような知見はインサイトと呼ばれる。一般的なデータ分析作業では、分析者が、仮説を設定し、設定した仮説に基づいてデータ分析・可視化し、その仮説検証する、というサイクルを繰り返すことによってインサイトを見出している。

[0003] インサイトを見出すための上記のようなデータ分析作業は、非常に時間と労力を要するものであるため、これを自動化する技術の開発が進められている。例えば、下記の特許文献1には、データセットから自動でインサイトを提供するシステムが開示されている。分析者は、特許文献1に記載のシステムに、分析したい多次元データを入力すればよい。これにより、当該システムにより自動的にインサイトが決定され、決定されたインサイトがディスプレイに表示される。

先行技術文献

特許文献

[0004] 特許文献1：米国特許第2020／0257682号明細書

発明の概要

発明が解決しようとする課題

[0005] 特許文献1に記載の技術には、複数のデータセット間のインサイトを検出することができないという点で改善の余地があった。例えば、ある企業の製品販売データからなるデータセットと、他の企業についての製品販売データからなるデータセットの両方を解析することにより、一方のデータセットの

みからは得られないインサイトが見つかる可能性がある。

[0006] しかしながら、特許文献 1 に記載の技術では、このような複数のデータセット間のインサイトを検出することは想定されていない。このため、当然のことながら、特許文献 1 に記載の技術では、複数のデータセット間のインサイトを検出することはできない。

[0007] 本発明の一態様は、上記の問題に鑑みてなされたものであり、その目的の一例は、複数のデータセット間におけるインサイトの検出を可能にする情報処理装置等を提供することである。

課題を解決するための手段

[0008] 本発明の一態様に係る情報処理装置は、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する分類手段と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する評価手段とを備える。

[0009] 本発明の一態様に係る分析方法は、少なくとも 1 つのプロセッサが、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化することと、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出すること、を含む。

[0010] 本発明の一態様に係る分析プログラムは、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する処理と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する処理と、をコンピュータに実行させる。

発明の効果

[0011] 本発明の一態様によれば、複数のデータセット間におけるインサイトの検出が可能になる。

図面の簡単な説明

[0012] [図1]本発明の例示的实施形態1に係る情報処理装置の構成を示すブロック図である。

[図2]本発明の例示的实施形態1に係る分析方法の流れを示すフロー図である。

[図3]本発明の例示的实施形態2に係る情報処理装置が実行する処理の概要を示す図である。

[図4]本発明の例示的实施形態2に係る情報処理装置の構成を示すブロック図である。

[図5]本発明の例示的实施形態2に係る分析方法の流れを示すフロー図である。

[図6]分析対象データと、当該分析対象データから生成されたインサイトサブジェクトの例を示す図である。

[図7]評価結果データと出力データの例を示す図である。

[図8]本発明の例示的实施形態3に係る情報処理装置の構成を示すブロック図である。

[図9]本発明の例示的实施形態3に係る分析方法の流れを示すフロー図である。

[図10]インサイトスコアの算出方法と、外れ値の検出方法を説明する図である。

[図11]上記情報処理装置の各機能を実現するソフトウェアであるプログラムの命令を実行するコンピュータの一例を示す図である。

発明を実施するための形態

[0013] [例示的实施形態1]

本発明の第1の例示的实施形態について、図面を参照して詳細に説明する。本例示的实施形態は、後述する例示的实施形態の基本となる形態である。

[0014] (情報処理装置 1 の構成)

本例示的实施形態に係る情報処理装置 1 の構成について、図 1 を参照して説明する。図 1 は、情報処理装置 1 の構成を示すブロック図である。図示のように、情報処理装置 1 は、分類部 1 1 と評価部 1 2 を備えている。

[0015] 分類部 1 1 は、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する。グループ化の際に、分類部 1 1 は、評価部 1 2 による評価値の算出が可能なインサイトサブジェクトをグループ化する。なお、以下では、検出対象のインサイトをインサイトタイプと呼ぶ。インサイトタイプは少なくとも 1 つ設定されていればよい。インサイトタイプの詳細は例示的实施形態 2 で説明する。

[0016] そして、評価部 1 2 は、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する。以下では、この評価値をインサイトスコアと呼ぶ。

[0017] 例えば、ある店舗の月間の売上記録を示すデータセットが分析対象である場合、その店舗における日別の総売上を示すデータ（日付と総売上のデータ項目を関連付けたデータ）をインサイトサブジェクトとすることができる。同様に、その店舗におけるある商品の日別の売上を示すデータ（日付とある商品の売上のデータ項目を関連付けたデータ）をインサイトサブジェクトとすることができる。このようなインサイトサブジェクトは、例えばチャート等の形式で可視化することができるため、インサイトサブジェクトを可視化パターンと呼ぶこともできる。インサイトサブジェクトは、多次元データであるデータセットから得られる各可視化パターンを特徴づけるものであると言うこともできる。この場合、1 つのインサイトサブジェクトにつき 1 つの可視化パターンが対応付けられる。

[0018] そして、検出対象のインサイト、すなわちインサイトタイプが、例えばインサイトサブジェクト間の相関であれば、分類部 1 1 は、相関の有無を判定

するためのインサイトスコア（例えば相関係数）の算出が可能なインサイトサブジェクトをグループ化する。例えば、分類部 11 は、上記の例では、各店舗における日付と売上の関係を示すインサイトサブジェクトをグループ化してもよい。これにより、評価部 12 は、各店舗における日付と売上についてインサイトスコアを算出することができる。インサイトスコアは、そのまま出力してもユーザがインサイトを発見する大きな助けとなる。また、インサイトスコアを用いることにより、インサイトスコアが高い、すなわちインサイトである可能性が高いインサイトサブジェクトの組み合わせを自動で検出することも可能になる。

[0019] 以上のように、本例示的实施形態に係る情報処理装置 1 では、複数のデータセットのそれぞれから生成されたインサイトサブジェクトを、検出対象のインサイトごとにグループ化する分類部 11 と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する評価部 12 と、を備える、という構成が採用されている。

[0020] したがって、本例示的实施形態に係る情報処理装置 1 によれば、複数のデータセット間におけるインサイトの検出が可能になるという効果が得られる。言い換えれば、本例示的实施形態に係る情報処理装置 1 によれば、複数のデータセットを横断的に分析することで得られる複合インサイト（以下、横断的複合インサイトと呼ぶ）の発見に繋がる可能性のあるデータをユーザに提示することが可能になる。

[0021] なお、上述の情報処理装置 1 の機能は、プログラムによって実現することもできる。本例示的实施形態に係る分析プログラムは、コンピュータに、複数のデータセットのそれぞれから生成されたインサイトサブジェクトを、検出対象のインサイトごとにグループ化する処理と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する処理と、を実行させる。したがって、本例示的实施形態に係る分析プログラムによれば、複数のデータセット間における

インサイト、すなわち横断的複合インサイトの検出が可能になるという効果が得られる。

[0022] (分析方法の流れ)

本例示的实施形態に係る分析方法の流れについて、図2を参照して説明する。図2は、本例示的实施形態に係る分析方法の流れを示すフロー図である。

[0023] S11では、少なくとも1つのプロセッサが、複数のデータセットのそれぞれから生成されたインサイトサブジェクトを、インサイトタイプごとにグループ化する。そして、S12では、少なくとも1つのプロセッサが、S11でグループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値であるインサイトスコアを算出する。これにより、図2の分析方法は終了する。

[0024] なお、1つのプロセッサにS11～S12の処理を実行させてもよいし、S11の処理とS12の処理をそれぞれ別のプロセッサに実行させてもよい。後者の場合、各プロセッサは、1つの情報処理装置が備えているものであってもよいし、それぞれ異なる情報処理装置が備えているものであってもよい。また、S11～S12の処理を実行する少なくとも1つのプロセッサは、情報処理装置1が備えているものであってもよい。

[0025] 以上のように、本例示的实施形態に係る分析方法においては、少なくとも1つのプロセッサが、複数のデータセットのそれぞれから生成されたインサイトサブジェクトをインサイトタイプごとにグループ化すること、およびグループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するためのインサイトスコアを算出すること、を含む、という構成が採用されている。このため、本例示的实施形態に係る分析方法によれば、複数のデータセット間におけるインサイト、すなわち横断的複合インサイトの検出が可能になるという効果が得られる。

[0026] [例示的实施形態2]

(概要)

本発明の第2の例示的实施形態について、図面を参照して詳細に説明する。本例示的实施形態では、複数のデータセットの入力を受け付けて、それらのデータセットについてのインサイトに関する情報を出力する情報処理装置2について説明する。図3は、情報処理装置2が実行する処理の概要を示す図である。

[0027] まず、情報処理装置2は、分析対象となる分析対象データ211aと211bを取得する。分析対象データ211aと211bは、何れも複数のレコードを含む多次元データのデータセットである。なお、分析対象データ211aと211bを区別する必要がないときには単に分析対象データ211と記載する。図3に示す分析対象データ211aと211bは何れもテーブル形式のデータである。

[0028] 次に、情報処理装置2は、取得した分析対象データ211aと211bのそれぞれからインサイトサブジェクトを生成する。図3の例では、分析対象データ211aから $l_1 \sim l_3$ の3つのインサイトサブジェクトが生成され、分析対象データ211bから l_4 、 l_5 の2つのインサイトサブジェクトが生成されている。

[0029] 続いて、情報処理装置2は、生成したインサイトサブジェクト $l_1 \sim l_5$ をグループ化する。図3の例では、インサイトサブジェクト l_1 と l_5 がグループ G^1 に分類され、インサイトサブジェクト l_3 と l_4 がグループ G^2 に分類されている。グループ G^1 と G^2 のインサイトタイプは同じであってもよいし、異なってもよい。ただし、グループ G^1 と G^2 のインサイトタイプが同じである場合には、各グループにはそれぞれ異なるインサイトサブジェクトを分類する。

[0030] そして、情報処理装置2は、各グループに含まれるインサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値であるインサイトスコアを算出する。図3の例では、インサイトサブジェクト l_1 と l_5 のインサイトスコアが0.6、インサイトサブジェクト l_3 と l_4 のインサイトスコアが0.9と算出されている。インサイトスコアは、例えばインサ

イトサブジェクト間の相関の程度を0～1の数値（数値が大きいほど相関の程度が高い）で示すものであってもよい。この場合、インサイトサブジェクト 1_3 と 1_4 は、相関が高いことになる。

[0031] ここで、インサイトサブジェクト 1_3 は、分析対象データ211aから生成されたものである。一方、インサイトサブジェクト 1_4 は、分析対象データ211bから生成されたものである。そして、インサイトサブジェクト 1_3 と 1_4 の相関が高いという知見は、人にとって有用なものである。つまり、情報処理装置2によれば、複数のデータセット間におけるインサイト、すなわち横断的複合インサイトの検出が可能になる。なお、詳細は以下説明するが、情報処理装置2は、相関以外にも様々なインサイトの検出を可能にする。

[0032] （情報処理装置2の構成）

図4は、情報処理装置2の構成を示すブロック図である。情報処理装置2は、情報処理装置2の各部を統括して制御する制御部20と、情報処理装置2が使用する各種データを記憶する記憶部21を備えている。また、情報処理装置2は、情報処理装置2が他の装置と通信するための通信部22、情報処理装置2に対する入力を受け付ける入力部23、および情報処理装置2がデータを出力するための出力部24を備えている。以下では、出力部24がデータを表示出力する表示装置である例を説明するが、出力部24の出力態様は任意であり、例えば印字出力や音声出力等の態様でデータを出力するものであってもよい。また、入力部23と出力部24は、情報処理装置2に外付けされた、情報処理装置2の外部の機器であってもよい。

[0033] 制御部20には、データ取得部201、サブジェクト生成部202、表記統一部203、分類部204、粒度統一部205、評価部206、および出力データ生成部207が含まれている。また、記憶部21には、分析対象データ211、評価結果データ212、および出力データ213が記憶されている。

[0034] 分析対象データ211は、情報処理装置2による分析対象の対象となるデータである。分析対象データ211には、複数のデータセットが含まれてい

る。各データセットは、複数のレコードを含む多次元データである。また、評価結果データ 212 は、評価部 206 による分析対象データ 211 の評価の結果を示すデータである。そして、出力データ 213 は、情報処理装置 2 による分析対象データ 211 の分析の結果をユーザに提示するためのデータ、すなわち分析対象データ 211 のインサイトに関するデータである。

[0035] データ取得部 201 は、情報処理装置 2 が分析する対象となる複数のデータセットを取得し、それらを分析対象データ 211 として記憶部 21 に記憶させる。データ取得部 201 は、分析開始時までに分析対象データ 211 を取得して記憶部 21 に記憶させればよい。分析対象データ 211 の取得方法は特に限定されない。例えば、データ取得部 201 は、情報処理装置 2 のユーザが入力部 23 を介して入力したデータセットを取得してもよい。また、例えば、データ取得部 201 は、通信部 22 を介した通信により、外部の装置から分析対象データ 211 を取得してもよい。

[0036] サブジェクト生成部 202 は、分析対象データ 211 に含まれる複数のデータセットのそれぞれからインサイトサブジェクトを生成する。より詳細には、サブジェクト生成部 202 は、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることによりインサイトサブジェクトを生成する。例えば、あるデータセットが、日付、売上、および場所のデータ項目を含む多次元データである場合、サブジェクト生成部 202 は、日付と売上を関連付けたインサイトサブジェクトや、場所と売上を関連付けたインサイトサブジェクトを生成する。

[0037] 表記統一部 203 は、各インサイトサブジェクトにおけるデータの表記を統一する。より詳細には、表記統一部 203 は、各インサイトサブジェクトに含まれる単語の中から類似した単語を抽出し、それらの単語を 1 つの単語に置き換えることにより、各インサイトサブジェクトにおける表記を統一する。なお、上記「類似」には、単語の文字列の類似の他、意味の類似も含まれる。

[0038] 例えば、あるデータセットにおいて商品の販売地を表す「東京都」は、他

のデータセットにおいて商品の販売地を表す「東京」と意味および文字列が類似した単語であり、これらは表記ゆれと呼ぶこともできる。また、例えば、あるデータセットにおいて商品の販売地を表す「都道府県」は、他のデータセットにおいて商品の販売地を表す「場所」と、意味が類似した単語である。

[0039] このような類似の単語を抽出する方法としては任意のものが適用可能である。表記統一部203は、「東京」と「東京都」のような表記ゆれの単語を抽出してもよい。この場合、表記統一部203は、例えば、単語間の編集距離が近い単語を抽出してもよい。編集距離は、レーベンシュタイン距離とも呼ばれ、2つの文字列がどの程度異なっているかを示す距離である。編集距離を求める際には、表記統一部203は、比較対象の一方の単語を構成する文字列に対して何回の変更処理（削除、挿入、置換）を行えば、比較対象の他方を構成する文字列に変換できるかを求める。この他にも、分析対象データ211は、例えば2つの文字列の長さとの置換の要不要（部分的な一致）を測る距離であるジャロ・ウィンクラー距離に基づいて類似の単語を抽出してもよい。

[0040] また、意味が類似した単語を抽出する場合、分析対象データ211は、例えば、各データセットに含まれる各単語を分散表現で表し、分散表現の類似度が高い単語を抽出してもよい。分散表現の導出には、例えばword2vec等のプログラムを用いることができる。

[0041] 表記統一部203は、類似した単語を抽出した後、それらの単語の表記を統一する。例えば、表記統一部203は、類似する2つの単語のうち一方の単語を他方の単語に全て置換することにより表記を統一してもよい。また、表記統一部203は、類似する2つの単語を、それらの単語を包括する上位概念的な単語に置換することにより表記を統一してもよい。

[0042] 分類部204は、サブジェクト生成部202が生成したインサイトサブジェクトをグループ化する。より詳細には、分類部204は、インサイトの有無を判定するための評価値であるインサイトスコアを算出可能なインサイト

サブジェクトをグループ化する。これにより、インサイトスコアに基づいてインサイトを検出することが可能になる。なお、1つのグループには任意の数のインサイトサブジェクトを含めることができる。そして、1つのグループには異なるデータセットから得られたインサイトサブジェクトを含めることができる。1つのグループには少なくとも1つのインサイトサブジェクトを含めることが好ましい。

[0043] なお、表記統一部203が複数のインサイトサブジェクトにおける表記を統一していた場合、評価部206は、表記が統一されたインサイトサブジェクトをグループ化する。異なるデータセット間では、表記が不統一であることも多く、表記が不統一であることが評価の支障となることも一般的には多いが、情報処理装置2によればそのような場合にも評価を行うことができる。つまり、情報処理装置2によれば、例示的实施形態1に係る情報処理装置1の奏する効果に加えて、表記が不統一なデータセットについても横断的複合インサイトを検出することが可能になるという効果が得られる。

[0044] 例えば、年別の売上を示すインサイトサブジェクトが複数存在する場合、それらのインサイトサブジェクトの系列名は何れも「年」と「売上」となるから、分類部204は、それらを1つのグループに分類する。また、このようなインサイトサブジェクトの一部で、系列名が「売上」等の他の表記となっていた場合でも、表記統一部203が表記を統一するため、分類部204は、それらを1つのグループに分類することができる。

[0045] ここで、上記のとおり、グループ化はインサイトタイプごとに行われる。よって、各インサイトタイプについて、グループ化の基準を予め定めておけばよい。インサイトタイプとしては、例えば相関が挙げられる。インサイトタイプが相関であるインサイトサブジェクトをグループ化する場合、分類部204は、相関関係の強さを評価できる、言い換えれば相関係数を計算可能なインサイトサブジェクトをグループ化すればよい。また、インサイトタイプが外れ値であるインサイトサブジェクトをグループ化する場合、分類部204は、外れ値を検出できるインサイトサブジェクト、つまり対応するデー

タ間の距離を計算可能なインサイトサブジェクトをグループ化すればよい。
具体的には、例えば、分類部204は、各系列名を示す単語が同一のインサイトサブジェクトを1つのグループに分類してもよい。

[0046] インサイトタイプとしては、相関以外にも任意のものを採用することができる。横断的複合インサイトを検出する場合、例えば、相互メジャー相関 (Cross-measure correlation)、二次元クラスタリング、帰属 (Attribution) 等のインサイトタイプを設定してもよい。

[0047] また、例えば、分類部204は、シングルポイントインサイト (Single point insight)、すなわち1つのインサイトサブジェクトを入力とする横軸に順序が存在しない (non-ordinal dimension) インサイトサブジェクトをグループ化してもよい。このようなグループ化により、例えば、突出したNo. 1 (Outstanding No.1)、突出した最下位 (Outstanding No. Last)、突出した上位2つ (Outstanding Top 2)、または均一度 (Evenness) 等のインサイトを検出することが可能になる。

[0048] また、分類部204は、シングルシェープインサイト (Single shape insight)、すなわち1つのインサイトサブジェクトを入力とする横軸に順序が存在する (ordinal dimension) インサイトサブジェクトをグループ化してもよい。なお、横軸に順序が存在するデータとしては例えば時系列データが挙げられる。このようなグループ化により、変化点 (Change point)、トレンド、季節性 (Seasonality)、外れ値等のインサイトを検出することが可能になる。設定されるインサイトタイプには、横断的複合インサイトを検出可能なもの (例えば相関等) が少なくとも1つ含まれていればよく、横断的ではない複合インサイトを検出するためのもの (例えば変化点 (Change point) 等) が含まれていてもよい。

[0049] 粒度統一部205は、各インサイトサブジェクトにおけるデータの粒度を統一する。この処理は、評価部206がインサイトサブジェクト間の関連性を評価できるようにするための処理であるから、粒度が揃っていないデータを対象として行われる。粒度の統一は、データセットから生成されたインサ

イトサブジェクトに対して行ってもよいし、分析対象となる複数のデータセットに対して予め行っておいてもよい。なお、データの粒度は、一連のデータがどのような細かさ（単位）であるかを示す。

[0050] 例えば、あるインサイトサブジェクトと他のインサイトサブジェクトが何れも月別の売上を示すものであるが、前者には毎月の売上が示されており、後者には隔月（奇数月）の売上が示されている場合、これらのデータの粒度は一致していない。この場合、両データ間の距離や類似度の評価ができないことがある。

[0051] 粒度統一部 205 は、このようなデータに対して粒度を揃える処理を行う。例えば、粒度統一部 205 は、欠損値補完によりデータを補完して粒度を揃えてもよいし、ダウンサンプリングにより粒度を揃えてもよい。欠損値補完は、他のデータから欠損部を予測して補完する処理であり、具体例としては内挿等が挙げられる。ダウンサンプリングは、サンプリング粒度を粗い方に合わせる処理である。

[0052] 上記の例において欠損値補完を行う場合、粒度統一部 205 は、他のインサイトサブジェクトにおける偶数月の売上を補完する。また、上記の例においてダウンサンプリングを行う場合、粒度統一部 205 は、あるインサイトサブジェクトにおける奇数月の売上のみが評価部 206 による評価に用いられるようにする。

[0053] 評価部 206 は、分類部 204 により同じグループに分類された複数のインサイトサブジェクトの組み合わせについてインサイトスコアを算出し、その算出結果を示す評価結果データ 212 を生成して記憶部 21 に記憶させる。例えば、評価部 206 は、同じグループに分類されたインサイトサブジェクトの組み合わせを入力としてインサイトスコアを返す関数 f_T を用いて上記の評価を行ってもよい。

[0054] f_T は、インサイトタイプ T ごとに予め定義される関数であり、検出したいインサイトを与えるインサイトサブジェクトが入力されると高い値になるように設計される。インサイトタイプ T に対応するインサイトグループを G_T と

すると、インサイトスコアは下記の式で表される。

$$[0055] \quad (\text{インサイトスコア}) = f_T(l_1, l_2, \dots, l_n \mid l_i \in G_T)$$

評価部206は、同じグループに分類された複数のインサイトサブジェクトを組にして、各組のインサイトスコアを算出してもよい。この場合、2つのインサイトサブジェクトを入力とする f_T を用いればよい。例えば、 $l_1 \sim l_3$ の3つのインサイトサブジェクトがグループ化されている場合、評価部206は、 l_1 と l_2 、 l_1 と l_3 、および l_2 と l_3 の各組をそれぞれ f_T に入力することにより、各組のインサイトスコアを算出する。

[0056] インサイトスコアの算出方法は、インサイトタイプに応じたものとする。例えば、組にしたインサイトサブジェクト間の線形な相関の程度を評価する場合、評価部206は、ピアソン相関係数を算出する f_T を用いてインサイトスコアを算出してもよい。この他にも、例えば、評価部206は、スピアマン順位相関係数やコサイン類似度、対応するデータ間のユークリッド距離やEMD (Earth Mover's distance) 等をインサイトスコアとして算出してもよい。

[0057] なお、粒度統一部205がインサイトサブジェクトのデータの粒度を統一していた場合、評価部206は、粒度が統一された複数のインサイトサブジェクトの組み合わせについてインサイトスコアを算出する。異なるデータセット間では、データの粒度が不統一であることも多く、粒度が不統一であることが評価の支障となることも一般的には多いが、情報処理装置2によればそのような場合にも評価を行うことができる。すなわち、情報処理装置2によれば、例示的实施形態1に係る情報処理装置1の奏する効果に加えて、粒度が不統一なデータを含むデータセットについても横断的複合インサイトを検出することが可能になるという効果が得られる。

[0058] 出力データ生成部207は、評価結果データ212を用いて出力データ213を生成する。出力データ生成部207は、情報処理装置2の必須の構成要素ではないが、出力データ生成部207を設けることにより、情報処理装置2による分析の結果をより認識しやすい態様でユーザに提示することが可

能になる。

[0059] (分析方法の流れ)

本例示的实施形態に係る分析方法の流れについて図5～図7を参照して説明する。図5は、分析方法の流れを示すフロー図である。また、図6は、分析対象データ211と、当該分析対象データ211から生成されたインサイトサブジェクトの例を示す図である。そして、図7は、評価結果データ212と出力データ213の例を示す図である。

[0060] S21では、データ取得部201が、複数のデータセットの入力を受け付けて、分析対象データ211として記憶部21に記憶させる。例えば、データ取得部201は、入力部23を介して、図6に示す分析対象データ211の入力を受け付ける。分析対象データ211には、コンビニエンスストアにおける都道府県別の各月の売上を示すデータセット(D^S)と、スーパーマーケットにおける都道府県別の各月の売上を示すデータセット(D^T)が含まれる。

[0061] S22では、サブジェクト生成部202が、分析対象データ211に含まれる各データセットからインサイトサブジェクトを生成する。例えば、図6に示すデータセット D^S 、 D^T を用いる場合、サブジェクト生成部202は、データセット D^S からインサイトサブジェクト I^S_1 と I^S_2 を生成し、データセット D^T からインサイトサブジェクト I^T_1 と I^T_2 を生成することができる。

[0062] インサイトサブジェクト I^S_1 は、コンビニエンスストアにおける都道府県別の売上を示すものであり、図6では、 I^S_1 を売上の棒グラフ（横軸が都道府県、縦軸が売上）として示している。また、インサイトサブジェクト I^S_2 は、コンビニエンスストアにおける月毎の売上を示すものであり、図6では、 I^S_2 を売上の折れ線グラフ（横軸が日付、縦軸が売上）として示している。

[0063] 同様に、インサイトサブジェクト I^T_1 は、スーパーマーケットにおける都道府県別の売上を示すものであり、図6では、 I^T_1 を売上の棒グラフ（横軸が都道府県、縦軸が売上）として示している。また、インサイトサブジェク

ト I^T_2 は、スーパーマーケットにおける月毎の売上を示すものであり、図 6 では、 I^T_2 を売上の折れ線グラフ（横軸が日付、縦軸が売上）として示している。

[0064] インサイトサブジェクト I は、例えば下記のようなデータ形式とすることもできる。

$I = \{\text{subspace, breakdown, measure, aggregation}\}$

上記 “subspace”（サブスペース）は、多次元データであるデータセットに含まれるレコードをどのようにフィルタしたかを示す。上記 “subspace” は、各チャートの凡例に対応する。例えば、図 6 の I^S_2 の折れ線グラフにおける “subspace” は「東京都」である。フィルタリングを行わないことは、“*” 等の記号で表せばよい。

[0065] 上記 “breakdown”（ブレイクダウン）は、多次元データであるデータセットを集計するキーとして使用されるカラムを示す。上記 “breakdown” は、各チャートの横軸に対応する。例えば、図 6 の I^S_2 の折れ線グラフにおける “breakdown” は「日付」である。

[0066] 上記 “measure”（メジャー）は、多次元データであるデータセットにおいて数値データとして使用されるカラムを示す。上記 “measure” は、各チャートの縦軸に対応する。例えば、図 6 の I^S_2 の折れ線グラフにおける “measure” は「売上」の数値データである。

[0067] 上記 “aggregation”（アグリゲーション）は、“breakdown” ごとにデータを集計する方法（例えば関数）を示す。上記 “aggregation” の例としては、合計、平均、最大値、最小値等が挙げられる。集計に用いられる関数が「合計」である場合、“aggregation” は省略してもよい。

[0068] 例えば、図 6 に示す I^S_2 であれば、 $I^S_2 = \{ \{*, \text{東京都}\}, \text{日付}, \text{売上} \}$ と表すことができる。S 2 2 では、サブジェクト生成部 2 0 2 は、分析対象データ 2 1 1 に含まれる各データセットからこのようなデータ形式のインサイトサブジェクトを生成してもよい。

[0069] S 2 3 では、表記統一部 2 0 3 が、S 2 2 で生成された各インサイトサブ

ジェクトにおけるデータの表記を統一する。例えば、図6に示す I^{S_1} 、 I^{S_2} 、 I^{T_1} 、 I^{T_2} の中では、 I^{S_1} における横軸のラベル「都道府県」と、 I^{T_1} における横軸のラベル「場所」の意味が類似している。また、 I^{S_1} の系列名「東京都」、「大阪府」、「神奈川県」は、 I^{T_1} の系列名「東京」、「大阪」、「神奈川」のそれぞれと意味および表記が類似している。表記統一部203は、このような単語を抽出し、それらの表記を統一する。例えば、表記統一部203は、 I^{S_1} における横軸のラベルを「場所」に置換し、系列名「東京都」、「大阪府」、「神奈川県」を、それぞれ「東京」、「大阪」、「神奈川」に置換してもよい。

[0070] S24では、分類部204が、S22で生成されたインサイトサブジェクトであって、S23で表記が統一されたインサイトサブジェクトをグループ化する。例えば、図6に示す I^{S_1} 、 I^{S_2} 、 I^{T_1} 、 I^{T_2} のうち、縦軸と横軸のラベルが共通するインサイトサブジェクトをグループ化するとする。この場合、分類部204は、縦軸のラベルが「売上」で横軸のラベルが「場所」である I^{S_1} と I^{T_1} をグループ化する。 I^{S_1} の「都道府県」は表記統一部203により「場所」に置換済みであるからこのようなグループ化が可能になっている。また、分類部204は、縦軸のラベルが「売上」で横軸のラベルが「日付」である I^{S_2} と I^{T_2} をグループ化する。

[0071] I^{S_1} と I^{T_1} を含むグループを G^1 、 I^{S_2} と I^{T_2} を含むグループを G^2 とすると、グループ化の結果は下記のように表される。

$$I^{S_1}, I^{T_1} \in G^1$$

$$I^{S_2}, I^{T_2} \in G^2$$

S25では、粒度統一部205が、S24でグループ化されたインサイトサブジェクトに含まれるデータの粒度を統一する。例えば、図6に示す I^{S_2} の「日付」は、奇数月の1日であるのに対し、 I^{T_2} の「日付」は毎月の1日である。粒度統一部205は、このように粒度に差異があるデータを抽出し、それらのデータの粒度を揃える処理を行う。例えば、粒度統一部205は、 I^{T_2} の「日付」のデータのうち、奇数月のデータを抽出（すなわちダウン

サンプリング)することにより、「日付」データの粒度を揃えてもよい。また、粒度統一部205は、 I^{S_2} の偶数月のデータを欠損値補完することにより、「日付」データの粒度を揃えてもよい。なお、欠損値補完は、データのサンプリング日付にずれがある場合にも有効である。例えば、粒度統一部205は、毎月1日のデータと、毎月15日のデータの粒度を揃える場合、毎月15日のデータを欠損値補完することにより、毎月1日のデータを生成してもよい。

[0072] S26では、評価部206が、S24でグループ化され、S25でデータの粒度が統一されたインサイトサブジェクトの組み合わせを評価し、評価結果を評価結果データ212として記憶部21に記憶させる。より詳細には、評価部206は、同じグループに含まれるインサイトサブジェクトを組にして、その組についてのインサイトスコアを算出する、という処理を各グループについて行う。

[0073] 例えば、評価部206は、 $f_T(I_i, I_j)$ の式で表されるスコア関数、すなわち評価対象とする2つのインサイトサブジェクトを入力とし、インサイトスコアを出力とする関数を用いてインサイトスコアを算出してもよい。このスコア関数を用いる場合、グループ G^1 のインサイトスコアは $f_T(I^{S_1}, I^{T_1})$ 、グループ G^2 のインサイトスコアは $f_T(I^{S_2}, I^{T_2})$ と表される。

[0074] 評価部206は、上述のような評価結果をリスト化することにより、例えば図7に示すような評価結果データ212を生成してもよい。図7に示す評価結果データ212は、インサイトサブジェクトの組み合わせと、その組み合わせについて算出されたインサイトスコアとを示すテーブル形式のデータである。また、図7に示す評価結果データ212には、インサイトスコアの順位を示す「ランク」と、「インサイトタイプ」についても示されている。このように、評価部206は、インサイトサブジェクトの組み合わせと、その組み合わせについて算出されたインサイトスコアに加えて、評価に関する各種情報を含む評価結果データ212を生成してもよい。

[0075] S27では、出力データ生成部207が、S26で生成された評価結果デ

ータ 2 1 2 を用いて出力データ 2 1 3 を生成し、出力部 2 4 に出力させる。
例えば、図 7 に示す評価結果データ 2 1 2 を用いる場合、出力データ生成部 2 0 7 は、インサイトスコア（ランク）が最も高いインサイトサブジェクトの組み合わせを示す出力データ 2 1 3 を生成し、出力部 2 4 に出力させる。
これにより、図 5 の処理は終了する。

[0076] 出力データ 2 1 3 は、インサイトをユーザが認識しやすいように、当該インサイトを可視化したものであってもよい。可視化方法は、インサイトタイプに応じて決定すればよい。例えば、出力データ生成部 2 0 7 は、インサイトタイプが「相関」である場合、インサイトに関する情報として相関関係を表すのに適したチャート（例えば二次元の散布図）を出力データ 2 1 3 として生成してもよい。

[0077] 図 7 の下側には、評価結果データ 2 1 2 に示されるインサイトサブジェクトの組み合わせのうち、最もインサイトスコアが高かった（つまり、ランクが 1 の）ものについてのインサイトに関する情報の例を示している。具体的には、図 7 に示されるインサイトに関する情報には、スーパーマーケットとコンビニエンスストアの売上の相関を示す散布図と、インサイトの詳細を示すインサイト情報とが含まれている。インサイト情報には、インサイトタイプとインサイトスコアその他、各インサイトサブジェクトの詳細とその元になったデータセットが示されている。このような情報を出力部 2 4 に出力させることにより、情報処理装置 2 のユーザに、スーパーマーケットとコンビニエンスストアの売上の推移に強い相関がある、というインサイトを容易に認識させることができる。

[0078] 無論、出力データ生成部 2 0 7 が生成する情報は、インサイトをユーザに認識させることができるようなものであればよく、図 7 の例に限られない。例えば、出力データ生成部 2 0 7 は、最もインサイトスコアが高かったインサイトサブジェクトの組み合わせについて、各インサイトサブジェクトのチャートを生成し、これを出力データ 2 1 3 としてもよい。

[0079] なお、分析結果をユーザに提示する際に、必ずしも新たな出力データ 2 1

3を生成する必要はない。例えば、評価部206が、図7に示す評価結果データ212の全部または一部を出力部24に出力させることにより、分析結果をユーザに提示してもよい。また、評価部206は、ランクが1となった各インサイトサブジェクトや、インサイトスコアが所定の閾値以上となった各インサイトサブジェクトを構成するデータを出力させてもよい。このように、分析結果を出力させる態様は任意であり、図7のような例に限定されない。また、分析結果の可視化方法をユーザに選択させてもよい。この場合、出力データ生成部207は、ユーザが選択した方法で分析結果を可視化する。

[0080] このように、情報処理装置2は、複数のデータセットの分析結果として、インサイトの発見に繋がる可能性のあるチャートやデータ等を出力することができる。これにより、人手でチャートを比較する必要がなくなる。また、最終的にはインサイトをユーザが検討する場合であっても、分析に役立つようなデータセットを容易に絞り込むことができる。よって、分析・可視化に要する時間を大幅に短縮することができる。

[0081] また、情報処理装置2を用いることにより、全ての分析をユーザが行う場合に生じる判断基準のブレが発生する余地もない。さらに、分析をユーザが行う場合に生じる見逃しのリスク等も低減することができる。また、大規模なデータセットが分析対象である場合、ユーザによる複合インサイトの発見は困難であるが、情報処理装置2によれば、複合インサイト（横断的複合インサイトも含む）の発見が容易になる。

[0082] なお、図5のフローチャートにおいて、S23の処理は、S24の処理よりも先に行えばよく、例えばS21とS22の間に行ってもよい。また、S25の処理は、S26の処理よりも先に行えばよく、例えばS21とS22の間に行ってもよい。

[0083] （粒度の違いへの対応の変形例）

評価部206は、データの粒度が異なる複数のインサイトサブジェクトの組み合わせについてもインサイトスコアを算出可能な評価方法により、イン

サイトサブジェクトを評価してもよい。これにより、例示的实施形態1に係る情報処理装置1の奏する効果に加えて、粒度が不統一なデータを含むデータセットについても横断的複合インサイトを検出することが可能になるという効果が得られる。また、この場合、粒度統一部205を省略することができるという効果も得られる。

[0084] 例えば、インサイトサブジェクトにおける横軸のデータに順序が存在する（ordinal dimensionである）場合には、評価部206は、DTW（Dynamic Time Warping：動的時間伸縮法）や関数データ解析によりインサイトスコアを算出してもよい。なお、順序が存在するデータの例としては、例えば時系列データ等が挙げられる。DTWでは、 $s = (s_1, \dots, s_n)$ と $t = (t_1, \dots, t_m)$ の要素間の距離を総当りで計算したコスト行列 W の端（1, 1）から端（ n , n ）の最短経路を動的計画法で求める。DTWによれば、サンプルサイズが異なるデータ間の距離や類似度を計算可能であり、そのような距離や類似度をインサイトスコアの計算に用いることができる。また、関数データ解析を用いる場合、評価部206は、各インサイトサブジェクトのレコードを表現する連続的な関数を導出し、その関数を介してインサイトサブジェクト間の距離や類似度を計算し、それらをインサイトスコアの計算に用いることができる。

[0085] [例示的实施形態3]

本発明の第3の例示的实施形態について、図面を参照して詳細に説明する。上述の例示的实施形態において、インサイトサブジェクトをグループ化したときに、3つ以上のインサイトサブジェクトが1つのグループに分類されることがあり得る。このような場合、上述したスコア関数 $f_T(l_i, l_j)$ では、3つ以上のインサイトサブジェクトをまとめて評価することはできない。また、3つ以上のインサイトサブジェクトをまとめて評価する方法については、特許文献1にも記載も示唆もされていない。

[0086] 本例示的实施形態では、3つ以上のインサイトサブジェクトをまとめて評価することが可能な評価方法について図8～図10に基づいて説明する。図

8は、本例示的实施形態に係る情報処理装置3の構成を示すブロック図である。図9は、本例示的实施形態に係る分析方法の流れを示すフロー図である。図10は、インサイトスコアの算出方法と、外れ値の検出方法を説明する図である。

[0087] (情報処理装置3の構成)

図8に示すように、情報処理装置3は、評価部31と外れ値検出部32を備えている。なお、外れ値を検出する必要がない場合には外れ値検出部32を省略してもよい。評価部31は、図1に示した評価部12および図4に示した評価部206と同様に、グループ化された複数のインサイトサブジェクトの組み合わせについてインサイトスコアを算出する。評価部31は、3つ以上のインサイトサブジェクトをまとめて評価することができる点、言い換えれば3つ以上のインサイトサブジェクトにおけるインサイトの有無を示す1つのインサイトスコアを算出できる点で、評価部12、206と相違している。

[0088] 具体的には、評価部31は、グループ化された複数のインサイトサブジェクトを主成分分析することにより求めた、各主成分の寄与度の偏りの程度に基づいて当該インサイトサブジェクトの組み合わせについてのインサイトスコアを算出する。主成分分析は、任意の数のインサイトサブジェクトを対象として行うことができる。このため、本例示的实施形態に係る情報処理装置3によれば、例示的实施形態1、2に係る情報処理装置1、2の奏する効果に加えて、3つ以上のインサイトサブジェクトをまとめて評価することが可能になるという効果が得られる。なお、評価方法の詳細およびこのような評価が可能である理由については、図9および図10に基づいて後述する。

[0089] 外れ値検出部32は、評価部31による主成分分析により求められた主成分を用いて、グループ化された複数のインサイトサブジェクトに含まれるデータを表すことにより、当該データに含まれる外れ値を検出する。このため、本例示的实施形態に係る情報処理装置3によれば、例示的实施形態1、2に係る情報処理装置1、2の奏する効果に加えて、評価のために行った主成

分分析の結果を利用した効率のよい外れ値検出ができるという効果が得られる。なお、外れ値検出方法の詳細およびこのような方法で外れ値を検出することが可能である理由については、図 9 および図 10 に基づいて後述する。

[0090] (情報処理装置 3 が実行する処理の流れ)

情報処理装置 3 が実行する処理の流れを図 9 に基づいて説明する。なお、図 9 の処理の前に、複数のインサイトサブジェクトがグループ化済であるとする。つまり、図 8 には示していないが、本例示的实施形態では、情報処理装置 3 が分類部 11 (例示的实施形態 1) または分類部 204 (例示的实施形態 2) に相当する構成を備えていることを想定している。なお、情報処理装置 3 は、情報処理装置 2 が備える各種構成 (例えば、データ取得部 201 やサブジェクト生成部 202 等) の一部または全部を備えていてもよい。

[0091] S31 では、評価部 31 が、インサイトサブジェクトのグループを評価する。より詳細には、まず、評価部 31 は、評価対象のグループに含まれる各インサイトサブジェクトにおける、主成分分析の対象とするデータを特定する。例えば、インサイトサブジェクトが $l = \{\text{subspace, breakdown, measure, aggregation}\}$ の形式で表されていた場合、評価部 31 は、各インサイトサブジェクトにおける “measure” の項目のデータを主成分分析の対象とすればよい。

[0092] 次に、評価部 31 は、主成分分析の対象として特定したデータについて主成分分析を行う。例えば、評価部 31 は、各インサイトサブジェクトにおける “measure” の項目のデータから多次元の相関行列を生成し、この相関行列を用いて主成分分析を行ってもよい。主成分分析により、固有値と固有ベクトルが算出される。

[0093] 続いて、評価部 31 は、算出された固有値を用いて、各主成分の寄与率を算出する。各主成分の寄与率はその軸方向 (固有ベクトル) における情報量とみなすことができるから、各主成分の寄与率の偏り度合いを調べることで、インサイトサブジェクト間の相関の強さを定量的に評価することができる。

[0094] 例えば、図10には、相関がないインサイトサブジェクトを主成分分析して算出された各主成分の寄与率を示す棒グラフ1001と、相関があるインサイトサブジェクトを主成分分析して算出された各主成分の寄与率を示す棒グラフ1002を示している。なお、図10において、PC1は第1主成分、PC2は第2主成分、PC3は第3主成分である。

[0095] 棒グラフ1001では、PC1～PC3の寄与率は概ね同程度であり、主成分間での偏り度合いは小さい。一方、棒グラフ1002では、PC1の寄与率が最も高く、PC2の寄与率はその半分程度であり、PC3の寄与率はかなり小さく、全体として偏り度合いが大きい。このように、インサイトサブジェクト間の相関の有無は、各主成分の寄与率の偏り度合いに明瞭に反映される。

[0096] したがって、各主成分の寄与率の偏り度合いを定量的に評価すれば、その評価結果をインサイトスコアとすることができる。例えば、第1主成分の寄与率をインサイトスコアとしてもよい。これは、図10に示されるように、各主成分の寄与率の偏り度合いが大きい場合（棒グラフ1002）には、小さい場合（棒グラフ1001）と比べて第1主成分PC1の寄与率が大きいためである。

[0097] また、図10に示されるように、各主成分の寄与率の偏り度合いが大きい場合（棒グラフ1002）には、PC1～PC3の中で寄与率が突出して高いもの（具体的にはPC1）が存在する。一方、各主成分の寄与率の偏り度合いが小さい場合（棒グラフ1001）には、寄与率が突出して高いものは存在しない。このため、例えば、各主成分の寄与率を入力とし、入力された寄与率の中に突出して高いものが含まれているほど高い値を出力するスコア関数を用いてインサイトスコアを算出することもできる。

[0098] なお、インサイトサブジェクト間の非線形な相関を検出したい場合には、評価部31は、通常の主成分分析のかわりに、任意のカーネルを用いたカーネル主成分分析を実行してもよい。また、レコードのサンプリング粒度の違いなどで相関行列が計算できない場合には、評価部31は、関数データ解析

を用いた関数主成分分析を実行してもよい。

[0099] S 3 2 では、外れ値検出部 3 2 が、グループ化された各インサイトサブジェクトに含まれる外れ値の検出を行う。例えば、S 3 1 で各インサイトサブジェクトにおける“measure”の項目のデータを用いた評価が行われていた場合、外れ値検出部 3 2 も各インサイトサブジェクトにおける“measure”の項目のデータにおける外れ値を検出する。

[0100] 外れ値の検出は、S 3 1 における評価のために行われた主成分分析により求められた主成分を用いて、グループ化された複数のインサイトサブジェクトに含まれるデータを表すことにより行われる。

[0101] 図 1 0 の 1 0 0 3 は、サンプルデータを主成分分析して求めた第 1 主成分 P C 1 と第 2 主成分 P C 2 により当該サンプルデータを表した点を、縦軸を P C 2、横軸を P C 1 とする座標平面上にプロットしたものである。主成分分析後のプロットにおいて、他のデータと離れているデータは、元のサンプルデータにおいても他のデータと離れている。よって、1 0 0 3 において「外れ値」とされているプロットのように、他のデータから離れたデータを外れ値として検出すればよい。

[0102] 例えば、外れ値検出部 3 2 は、主成分で表されたデータの Hotelling の T^2 統計量を算出し、算出した T^2 統計量が顕著なデータを外れ値として検出してもよい。図 1 0 の 1 0 0 4 は、同図の 1 0 0 3 に示すサンプルデータから算出した T^2 統計量を、横軸がサンプル番号、縦軸が T^2 統計量の座標平面にプロットしたものである。同図の 1 0 0 3 において「外れ値」とされていたプロットは、 T^2 統計量が他のプロットと比べて大きい値となっている。よって、外れ値検出部 3 2 は、 T^2 統計量を用いて外れ値を検出することができる。

[0103] また、 T^2 統計量は F 分布や χ^2 分布に従うことが知られている。このため、外れ値検出部 3 2 は、統計的検定に基づいて得られた p 値を用いてスコアを計算してもよい。この場合、外れ値検出部 3 2 は、算出したスコアを用いて外れ値を検出すればよい。

[0104] 以上により、図 9 の処理は終了する。なお、S 3 1 の評価結果と S 3 2 で

検出された外れ値は、評価結果データとして記憶しておけばよい。評価結果データは、そのまま出力してもよいし、例示的实施形態2と同様に、評価結果データから出力データを生成し、生成した出力データを出力してもよい。

[0105] 〔参考例〕

評価部31による上述の評価方法は、横断的複合インサイトの検出に好適であると共に、横断的ではない、つまり1つのデータセットにおけるインサイトの検出にも好適である。このため、上述の情報処理装置3は、必ずしも分類部204（例示的实施形態2）や、分類部11（例示的实施形態1）に相当する構成を備えている必要はない。

[0106] 本参考例に係る情報処理装置3は、評価対象となる複数のインサイトサブジェクトを取得する取得部と、上述の評価部31を備えている。前記取得部が取得する複数のインサイトサブジェクトは、少なくとも1つのデータセットから生成されたものであればよい。つまり、複数のデータセットから生成された複数のインサイトサブジェクトを用いることが必須ではない点で、本参考例と上述の各例示的实施形態は相違している。

[0107] 本参考例の情報処理装置によれば、評価部31は、取得部が取得した複数の前記インサイトサブジェクトを主成分分析することにより得られた、各主成分の寄与度の偏りの程度に基づいて、当該インサイトサブジェクトの組み合わせについてのインサイトスコアを算出する。よって、3つ以上のインサイトサブジェクトをまとめて評価することができなかったという従来の課題を解決することができる。

[0108] また、本参考例に係る分析方法は、少なくとも1つのプロセッサが、評価対象となる複数のインサイトサブジェクトを取得すること、および、取得した複数の前記インサイトサブジェクトを主成分分析することにより得られた、各主成分の寄与度の偏りの程度に基づいて、当該インサイトサブジェクトの組み合わせについてのインサイトスコアを算出すること、を含む。そして、本参考例に係る分析プログラムは、コンピュータに、評価対象となる複数のインサイトサブジェクトを取得する処理と、取得した複数の前記インサイ

トサブジェクトを主成分分析することにより得られた、各主成分の寄与度の偏りの程度に基づいて、当該インサイトサブジェクトの組み合わせについてのインサイトスコアを算出する処理と、を実行させる。これらの分析方法および分析プログラムによっても、3つ以上のインサイトサブジェクトをまとめて評価することができなかったという従来の課題を解決することができる。

[0109] 〔変形例〕

上述の例示的实施形態1において、1つの情報処理装置1が行っていた処理は、複数の情報処理装置に分担させてもよい。言い換えれば、情報処理装置1が行う処理の一部を、少なくとも1つの他の情報処理装置に実行させてもよい。さらに言い換えれば、上述の各処理を少なくとも1つのプロセッサに行わせる場合、その少なくとも1つのプロセッサは、1つの情報処理装置1が備えているものであってもよいし、それぞれ異なる情報処理装置が備えているものであってもよい。これは、上述の例示的实施形態2における情報処理装置2、および例示的实施形態3における情報処理装置3についても同様である。

[0110] 〔ソフトウェアによる実現例〕

情報処理装置1～3の一部又は全部の機能は、集積回路（ICチップ）等のハードウェアによって実現してもよいし、ソフトウェアによって実現してもよい。

[0111] 後者の場合、情報処理装置1～3は、例えば、各機能を実現するソフトウェアであるプログラムの命令を実行するコンピュータによって実現される。このようなコンピュータの一例（以下、コンピュータCと記載する）を図11に示す。コンピュータCは、少なくとも1つのプロセッサC1と、少なくとも1つのメモリC2と、を備えている。メモリC2には、コンピュータCを情報処理装置1～3として動作させるためのプログラムPが記録されている。コンピュータCにおいて、プロセッサC1は、プログラムPをメモリC2から読み取って実行することにより、情報処理装置1～3の各機能が実現

される。

[0112] プロセッサC1としては、例えば、CPU (Central Processing Unit)、GPU (Graphic Processing Unit)、DSP (Digital Signal Processor)、MPU (Micro Processing Unit)、FPU (Floating point number Processing Unit)、PPU (Physics Processing Unit)、マイクロコントローラ、又は、これらの組み合わせなどを用いることができる。メモリC2としては、例えば、フラッシュメモリ、HDD (Hard Disk Drive)、SSD (Solid State Drive)、又は、これらの組み合わせなどを用いることができる。

[0113] なお、コンピュータCは、プログラムPを実行時に展開したり、各種データを一時的に記憶したりするためのRAM (Random Access Memory) を更に備えていてもよい。また、コンピュータCは、他の装置との間でデータを送受信するための通信インタフェースを更に備えていてもよい。また、コンピュータCは、キーボードやマウス、ディスプレイやプリンタなどの入出力機器を接続するための入出力インタフェースを更に備えていてもよい。

[0114] また、プログラムPは、コンピュータCが読み取り可能な、一時的でない有形の記録媒体Mに記録することができる。このような記録媒体Mとしては、例えば、テープ、ディスク、カード、半導体メモリ、又はプログラマブルな論理回路などを用いることができる。コンピュータCは、このような記録媒体Mを介してプログラムPを取得することができる。また、プログラムPは、伝送媒体を介して伝送することができる。このような伝送媒体としては、例えば、通信ネットワーク、又は放送波などを用いることができる。コンピュータCは、このような伝送媒体を介してプログラムPを取得することもできる。

[0115] [付記事項1]

本発明は、上述した実施形態に限定されるものでなく、請求項に示した範囲で種々の変更が可能である。例えば、上述した実施形態に開示された技術的手段を適宜組み合わせて得られる実施形態についても、本発明の技術的範囲に含まれる。

[0116] 〔付記事項 2〕

上述した実施形態の一部又は全部は、以下のようにも記載され得る。ただし、本発明は、以下の記載する態様に限定されるものではない。

[0117] （付記 1）

複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する分類手段と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する評価手段と、を備える情報処理装置。この構成によれば、複数のデータセット間におけるインサイトの検出を可能にすることができる。

[0118] （付記 2）

複数の前記インサイトサブジェクトにおける表記を統一する表記統一手段をさらに備え、前記分類手段は、表記が統一された前記インサイトサブジェクトをグループ化する、付記 1 に記載の情報処理装置。この構成によれば、表記が不統一なデータセットについても横断的複合インサイトを検出することが可能になる。

[0119] （付記 3）

複数の前記インサイトサブジェクトにおけるデータの粒度を統一する粒度統一手段をさらに備え、前記評価手段は、粒度が統一された複数の前記インサイトサブジェクトについて前記評価値を算出する、付記 1 または 2 に記載の情報処理装置。この構成によれば、粒度が不統一なデータを含むデータセットについても横断的複合インサイトを検出することが可能になる。

[0120] （付記 4）

前記評価手段は、動的時間伸縮法または関数データ解析により前記評価値を算出する、付記 1 または 2 に記載の情報処理装置。この構成によれば、粒度が不統一なデータを含むデータセットについても横断的複合インサイトを検出することが可能になる。

[0121] (付記5)

前記評価手段は、グループ化された複数の前記インサイトサブジェクトを主成分分析することにより求めた、各主成分の寄与度の偏りの程度に基づいて前記評価値を算出する、付記1から4の何れかに記載の情報処理装置。この構成によれば、3つ以上のインサイトサブジェクトをまとめて評価することが可能になる。

[0122] (付記6)

前記主成分分析により求められた主成分を用いて、グループ化された複数の前記インサイトサブジェクトに含まれるデータを表すことにより、当該データに含まれる外れ値を検出する外れ値検出手段をさらに備える、付記5に記載の情報処理装置。この構成によれば、評価のために行った主成分分析の結果を利用した効率のよい外れ値検出ができる。

[0123] (付記7)

少なくとも1つのプロセッサが、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化すること、およびグループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出すること、を含む分析方法。この構成によれば、複数のデータセット間におけるインサイトの検出を可能にすることができる。

[0124] (付記8)

コンピュータに、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する処理と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する処理と、を実行させる分析プログラム。この構成によれば、複数のデータセット間におけるインサイトの検出を可能にすることができる。

[0125] (付記9)

少なくとも1つのプロセッサを備え、前記プロセッサは、複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する処理と、グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する処理とを実行する情報処理装置。

[0126] なお、この情報処理装置は、更にメモリを備えていてもよく、このメモリには、前記をグループ化する処理と、前記評価する処理とを前記プロセッサに実行させるためのプログラムが記憶されていてもよい。また、このプログラムは、コンピュータ読み取り可能な一時的でない有形の記録媒体に記録されていてもよい。

符号の説明

- [0127]
- | | |
|-------|-----------------|
| 1 | 情報処理装置 |
| 1 1 | 分類部（分類手段） |
| 1 2 | 評価部（評価手段） |
| 2 | 情報処理装置 |
| 2 0 3 | 表記統一部（表記統一手段） |
| 2 0 4 | 分類部（分類手段） |
| 2 0 5 | 粒度統一部（粒度統一手段） |
| 2 0 6 | 評価部（評価手段） |
| 3 | 情報処理装置 |
| 3 1 | 評価部（評価手段） |
| 3 2 | 外れ値検出部（外れ値検出手段） |

請求の範囲

- [請求項1] 複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化する分類手段と、
- グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出する評価手段と、を備える情報処理装置。
- [請求項2] 複数の前記インサイトサブジェクトにおける表記を統一する表記統一手段を備え、
- 前記分類手段は、表記が統一された前記インサイトサブジェクトをグループ化する、請求項1に記載の情報処理装置。
- [請求項3] 複数の前記インサイトサブジェクトにおけるデータの粒度を統一する粒度統一手段を備え、
- 前記評価手段は、粒度が統一された複数の前記インサイトサブジェクトについて前記評価値を算出する、請求項1または2に記載の情報処理装置。
- [請求項4] 前記評価手段は、動的時間伸縮法または関数データ解析により前記評価値を算出する、請求項1または2に記載の情報処理装置。
- [請求項5] 前記評価手段は、グループ化された複数の前記インサイトサブジェクトを主成分分析することにより求めた、各主成分の寄与度の偏りの程度に基づいて前記評価値を算出する、請求項1から4の何れか1項に記載の情報処理装置。
- [請求項6] 前記主成分分析により求められた主成分を用いて、グループ化された複数の前記インサイトサブジェクトに含まれるデータを表すことにより、当該データに含まれる外れ値を検出する外れ値検出手段を備える、請求項5に記載の情報処理装置。
- [請求項7] 少なくとも1つのプロセッサが、

複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化すること、および

グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出すること、を含む分析方法。

[請求項8]

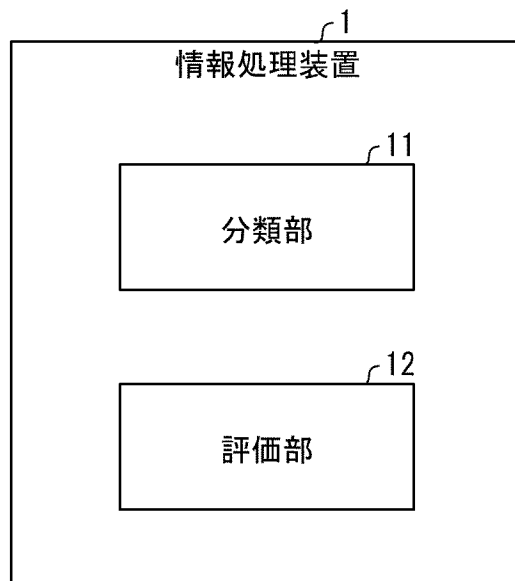
コンピュータに、

複数のデータセットのそれぞれから当該データセットに含まれる複数のデータ項目を関連付けることにより生成されたデータであるインサイトサブジェクトを、検出対象のインサイトごとにグループ化処理と、

グループ化された複数の前記インサイトサブジェクトの組み合わせについて、インサイトの有無を判定するための評価値を算出処理と、を実行させる分析プログラム。

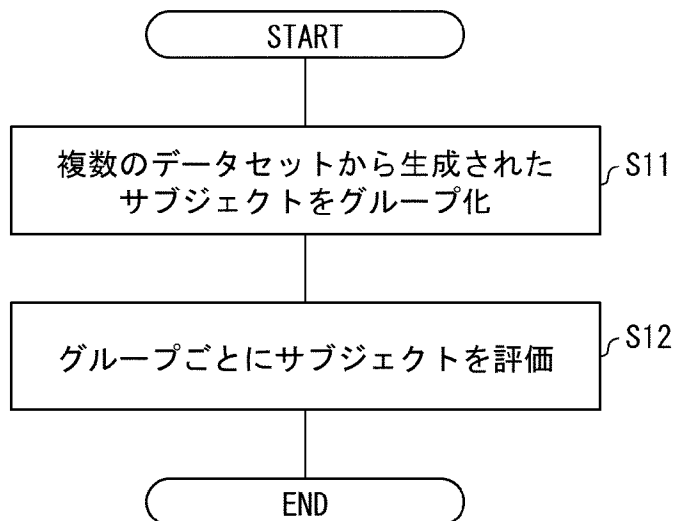
[図1]

図 1



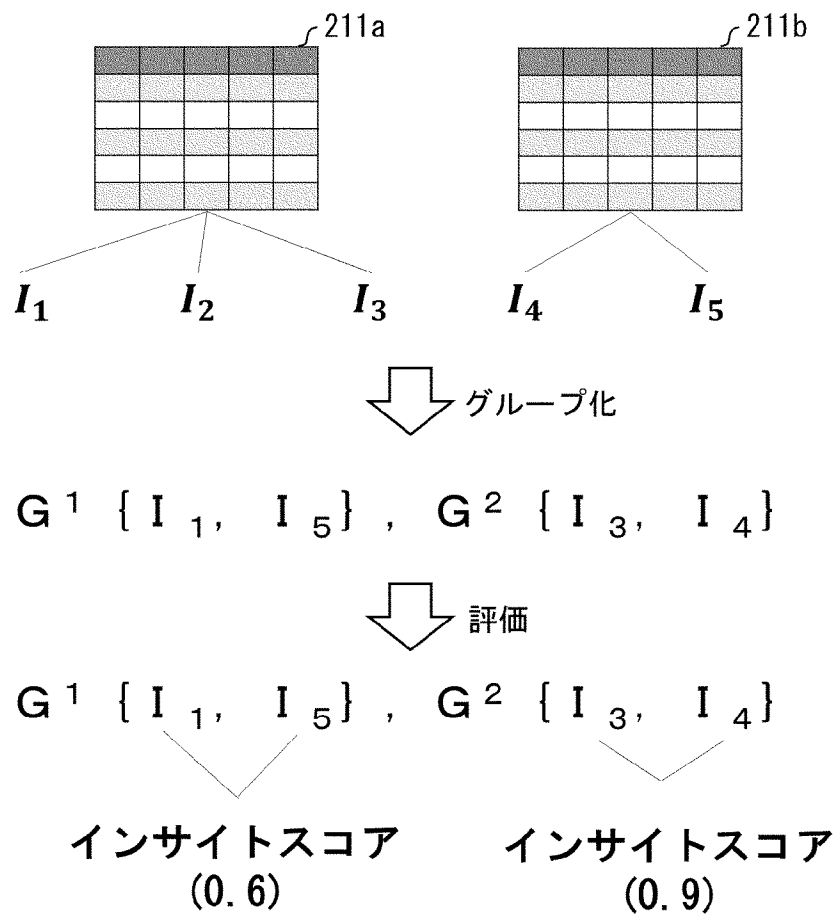
[図2]

図 2



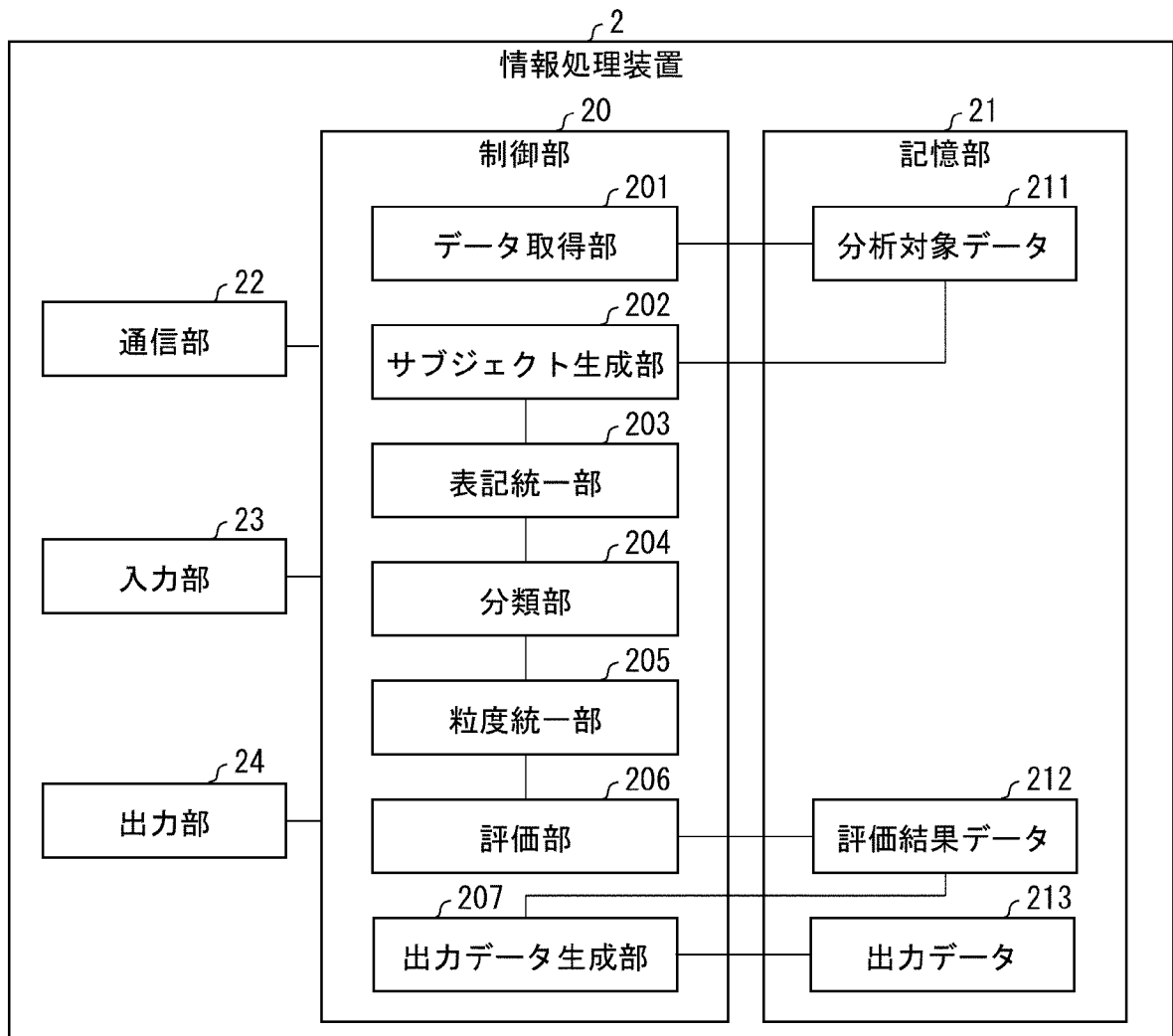
[図3]

図 3



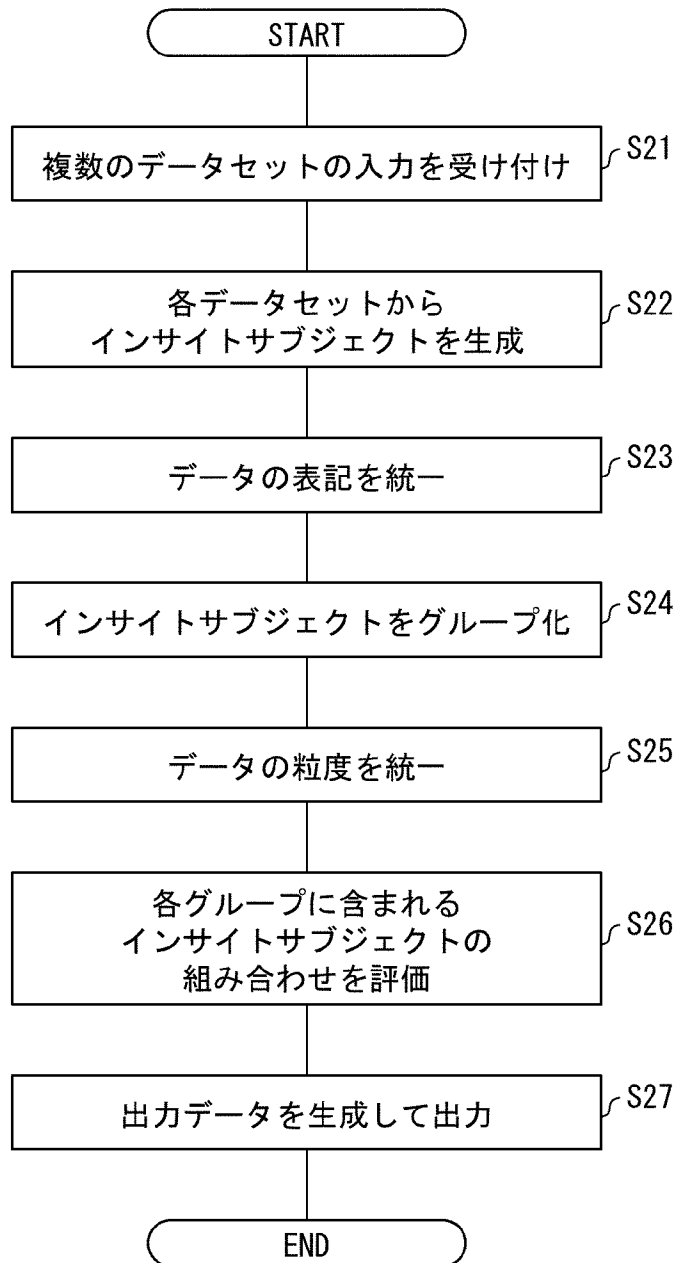
[図4]

図 4



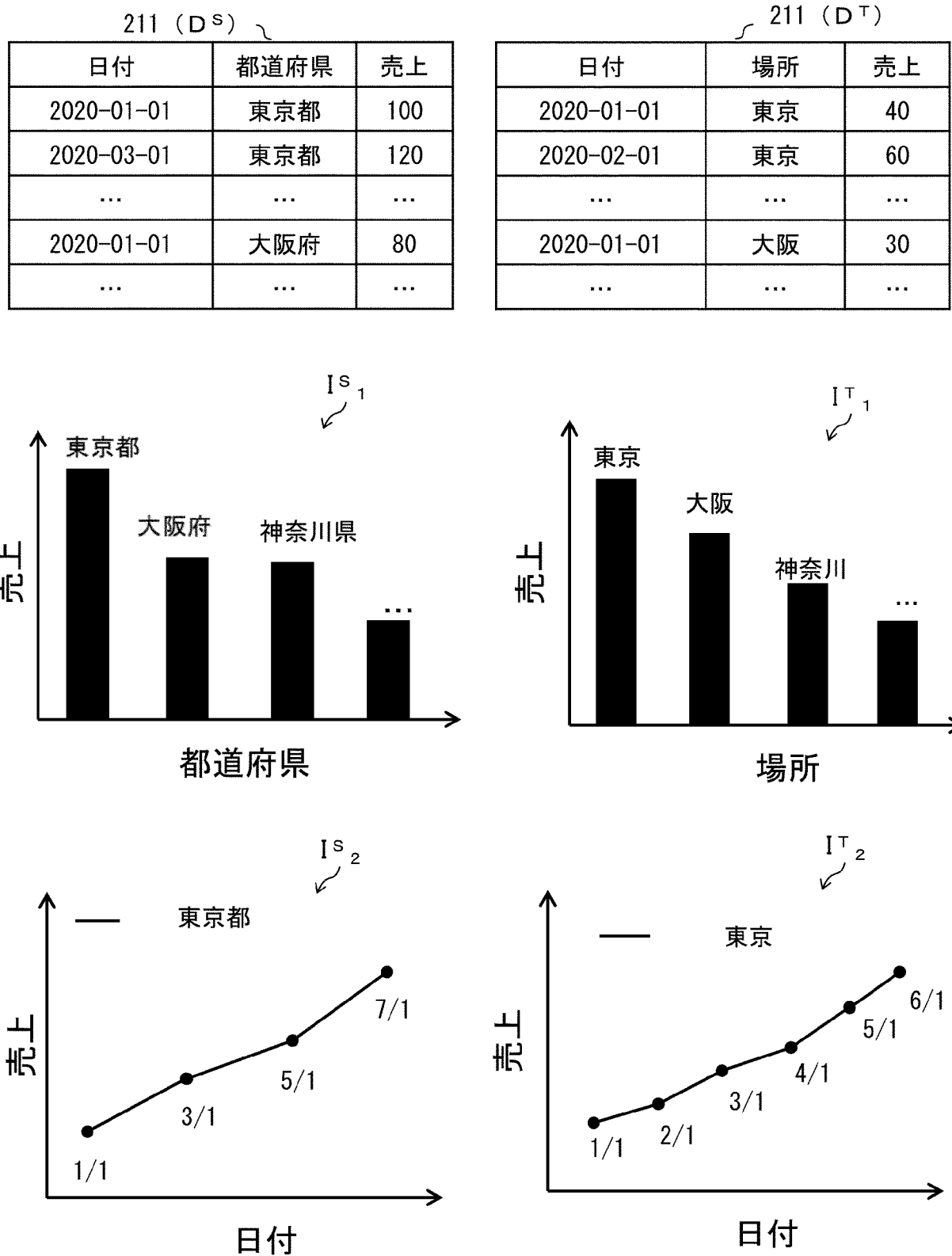
[図5]

図 5



[図6]

図 6



[図7]

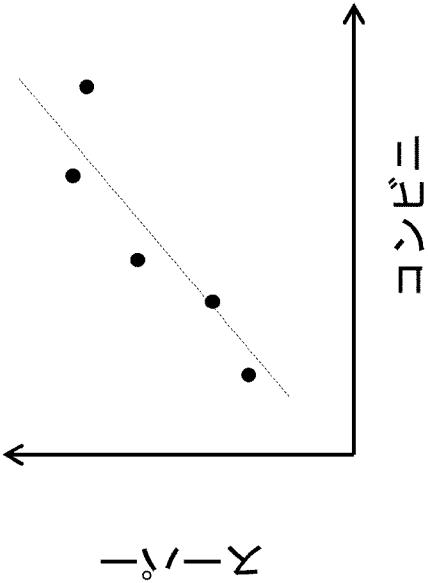
図 7

ランク	インサイトサブジェクト	インサイトスコア	インサイトタイプ
1	I^S_2, I^T_2	0.9	相関
2	I^S_1, I^T_1	0.3	相関
3	...	...	...

213

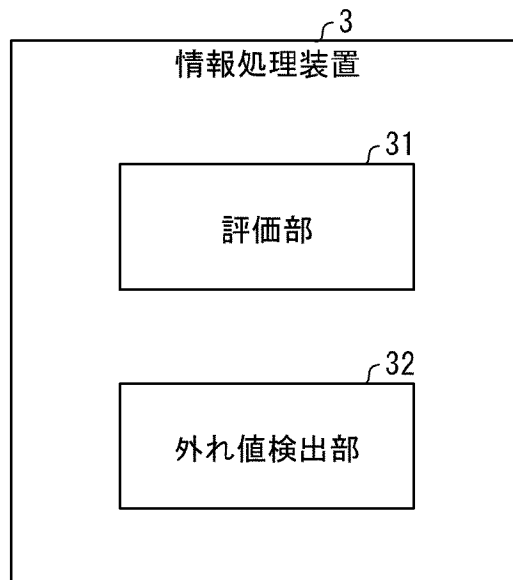
インサイト情報

- インサイトタイプ: 相関
 - インサイトスコア : 0.9
- データ 1
- データセット D^S ,
 - インサイトサブジェクト = $\{ \{*,*\}, \text{日付}, \text{売上} \} \}$
- データ 2
- データセット D^T
 - インサイトサブジェクト = $\{ \{*,*\}, \text{日付}, \text{売上} \} \}$



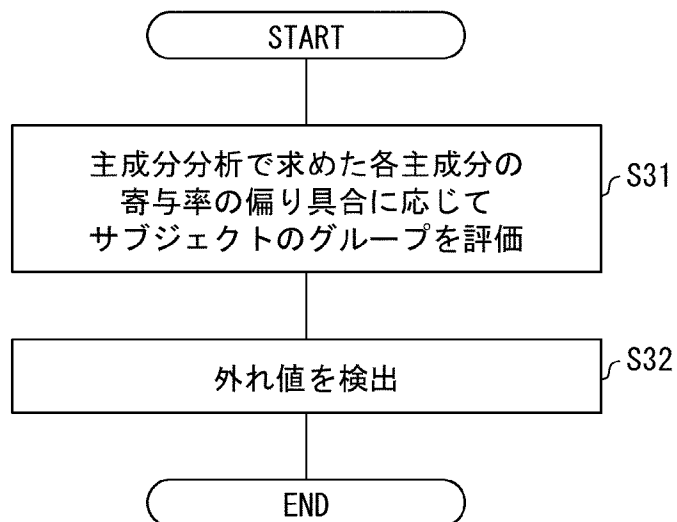
[図8]

図 8



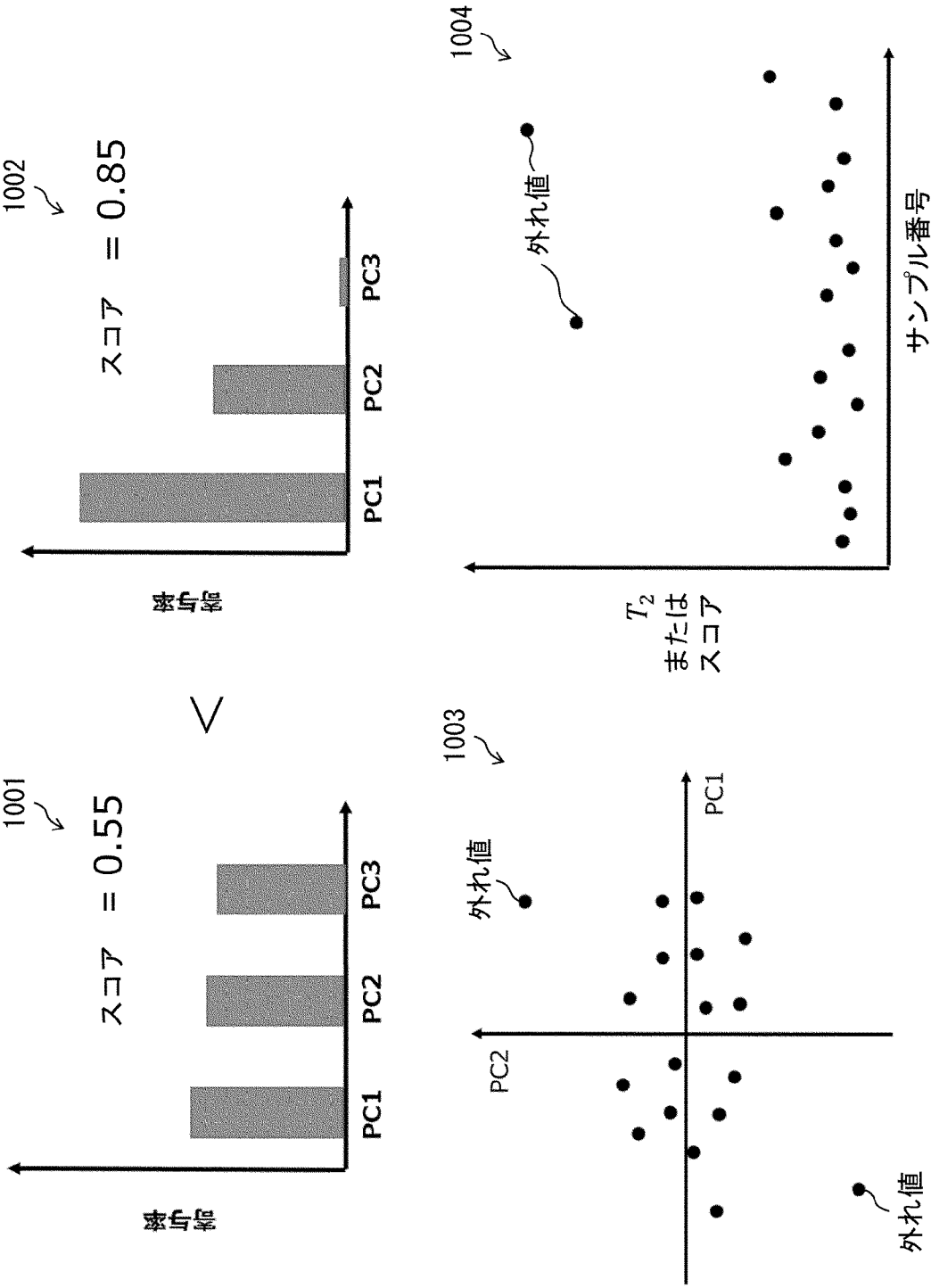
[図9]

図 9



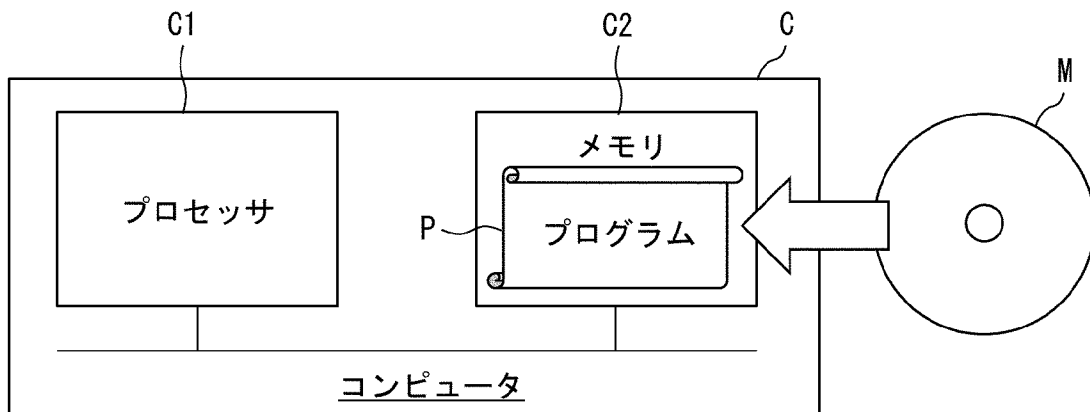
[図10]

図 10



[図11]

図 11



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/039367

A. CLASSIFICATION OF SUBJECT MATTER**G06Q 10/06**(2012.01)i

FI: G06Q10/06 300

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q10/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2022

Registered utility model specifications of Japan 1996-2022

Published registered utility model applications of Japan 1994-2022

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2017/163277 A1 (NEC CORP) 28 September 2017 (2017-09-28) claims 1-3, paragraphs [0019], [0020], [0028]-[0036], [0048]-[0053], [0085]-[0092]	1-5, 7, 8
Y	塚越 雄登, 次元間の関係に着目したドメインオントロジーに基づく異種データ間の関連性発見, 情報処理学会 研究報告 知能システム (ICS), 07 September 2020, ISSN:2188-885X pp. 1-8, (TSUKAGOSHI, Yuto. IPSJ Technical Report: Intelligent Computing Systems (ICS).), non-official translation (Finding relationships between heterogeneous data based on domain ontology focusing on relationships between dimensions)	1-5, 7, 8
Y	NEC、AIで予測分析した結果を可視化し、次の一手を提示するdotDataの新サービスを販売開始, [online], 07 October 2020, [retrieved on 06 January 2022], Internet <URL:https://web.archive.org/web/20201124042138/https://jpn.nec.com/press/202010/20201007_01.html> pp. 1-3, non-official translation (Starting sales of a new service of dotData that visualizes the results predicted and analyzed with AI and presents the next move)	1-5, 7, 8
A	US 2020/0257682 A1 (MICROSOFT TECHNOLOGY LICENSING, LLC) 13 August 2020 (2020-08-13) paragraphs [0065]-[0103]	1-8

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

11 January 2022

Date of mailing of the international search report

25 January 2022

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/039367**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2019-148897 A (HITACHI LTD) 05 September 2019 (2019-09-05) paragraphs [0046]-[0059]	1-8
P, A	JP 2021-43899 A (DAINIPPON PRINTING CO LTD) 18 March 2021 (2021-03-18) claims 1, 2, 5, 8	1-8

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2021/039367

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
WO	2017/163277	A1	28 September 2017	US 2019/0034945 A1 claims 1-3, paragraphs [0050], [0051], [0059]-[0069], [0080]- [0086], [0119]-[0126]	
US	2020/0257682	A1	13 August 2020	EP 3314462 A1 CN 107077489 A	
JP	2019-148897	A	05 September 2019	US 2019/0266619 A1 paragraphs [0058]-[0074]	
JP	2021-43899	A	18 March 2021	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06Q 10/06(2012.01)i FI: G06Q10/06 300		
B. 調査を行った分野		
調査を行った最小限資料（国際特許分類（IPC）） G06Q10/06		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2022年 日本国実用新案登録公報 1996-2022年 日本国登録実用新案公報 1994-2022年		
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	WO 2017/163277 A1（日本電気株式会社）28.09.2017（2017-09-28） 請求項1-3, 段落[0019], [0020], [0028]-[0036], [0048]-[0053], [0085]-[0092]	1-5, 7, 8
Y	塚越 雄登, 次元間の関係に着目したドメインオントロジーに基づく異種データ間の 関連性発見, 情報処理学会 研究報告 知能システム（ICS）, 2020.09.07, ISSN:2188-885X pp.1-8	1-5, 7, 8
Y	NEC, AIで予測分析した結果を可視化し、次の一手を提示するdotDataの新サー ビスを販売開始, [online], 2020.10.07, [2022年1月6日検索], インターネッ ト<URL:https://web.archive.org/web/20201124042138/https://jpn.nec.com/ press/202010/20201007_01.html> pp.1-3	1-5, 7, 8
A	US 2020/0257682 A1（MICROSOFT TECHNOLOGY LICENSING, LLC）13.08.2020（2020- 08-13） paragraphs[0065]-[0103]	1-8
A	JP 2019-148897 A（株式会社日立製作所）05.09.2019（2019-09-05） 段落[0046]-[0059]	1-8
☒ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的技術水準を示すもの “E” 国際出願日前の出願または特許であるが、国際出願日以後に 公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若し くは他の特別な理由を確立するために引用する文献（理由を 付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の 後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵 触するものではなく、発明の原理又は理論の理解のために引 用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性 又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文 献との、当業者にとって自明である組合せによって進歩性がな いと考えられるもの “&” 同一パテントファミリー文献		
国際調査を完了した日 11.01.2022		国際調査報告の発送日 25.01.2022
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号		権限のある職員（特許庁審査官） 山崎 誠也 5L 3978 電話番号 03-3581-1101 内線 3562

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
P, A	JP 2021-43899 A (大日本印刷株式会社) 18.03.2021 (2021 - 03 - 18) 請求項1, 2, 5, 8	1-8

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2021/039367

引用文献			公表日	パテントファミリー文献	公表日
WO	2017/163277	A1	28.09.2017	US 2019/0034945 A1 claims 1-3, paragraphs[0050], [0051], [0059]-[0069], [0080]-[0086], [0119]-[0126]	
US	2020/0257682	A1	13.08.2020	EP 3314462 A1 CN 107077489 A	
JP	2019-148897	A	05.09.2019	US 2019/0266619 A1 paragraphs[0058]-[0074]	
JP	2021-43899	A	18.03.2021	(ファミリーなし)	